International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

xx

COPY RIGHT

IJIEMR Transactions, online available on 19th JULY 2024. Link

https://www.ijiemr.org/downloads/Volume-13/Issue-7

10.48047/IJIEMR/V13/ISSUE 07/08

TITLE: **ADVANCED CLASSIFICATION OF EUROPEAN COURT OF HUMAN RIGHTS CASES THROUGH HYBRID TRANSFORMER-BASED TECHNIQUES**

Volume 13, ISSUE 07, Pages: 63-67

Paper Authors : Ravinder Reddy Baireddy , Dr. Rohita Y , Amireddy Poojitha , K. Nayana Sharon, N. Charan

USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper As Per UGC Guidelines We Are Providing A Electronic Bar Code

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

ADVANCED CLASSIFICATION OF EUROPEAN COURT OF HUMAN RIGHTS CASES THROUGH HYBRID TRANSFORMER-BASED TECHNIQUES

Ravinder Reddy Baireddy[1] , Dr. Rohita Y[2] , Amireddy Poojitha[3] , K. Nayana Sharon[4], N. Charan[5]

[1]Assistant Professor, IT Dept, SNIST, Hyderabad, India ravibaireddy@gmail.com.com

[2]Assoc Professor, IT Dept, SNIST, Hyderabad, India rohitay@sreenidhi.edu.in

[3, 4, 5] Student, IT Dept, SNIST, Hyderabad, India

20311A12M6@ sreenidhi.edu.in, 20311A12K5@ sreenidhi.edu.in, 20311A12J6@ sreenidhi.edu.in

ABSTRACT. In the realm of text classification, transformer-based models like BERT and its variants have gained significant recognition for their effectiveness. However, their application to classifying lengthy documents across various domains has been constrained by memory and computational limitations. Additionally, these models are often pre-trained on general language corpora, limiting their efficacy in domain-specific contexts such as legal documents. This study addresses the classification of legal documents, particularly from the European Court of Human Rights (ECHR) dataset, emphasizing the importance of efficient legal representation and the need for automated classification to enhance the legal system's efficiency and reduce costs. The proposed approach employs a sliding window technique to manage long documents and integrates traditional machine learning techniques with transformer-based models for classification. By leveraging methods such as Transfer Learning with BERT, RoBERTa, BigBird, Electra, and XLNet, the study aims to develop an accurate classifier for legal documentation, particularly focusing on human rights-related cases. As an extension, ensemble methods like the Voting Classifier, LSTM, and LSTM + GRU models are explored to further enhance classification accuracy, achieving a notable 92% accuracy rate. Additionally, a user-friendly frontend using the Flask framework with authentication support is proposed for user testing and accessibility. This research contributes to advancing automated legal document classification, facilitating improved legal aid provision, cost reduction, and enhanced accessibility to legal assistance.Keywords: *Legal documents classification, European court of human rights (ECHR) dataset, natural language processing, transformers, BERT, BigBird, ELECTRA, XLNet, legal-BERT.*

I. INTRODUCTION:

In contemporary society, the proliferation of legal rules and regulations underscores the paramount importance of legal representation and the defense of one's legal rights [1]. However, access to high-quality legal representation is often hindered by financial barriers, leading many individuals to rely on public legal aid programs for assistance [2]. Despite the existence of such programs, studies across various countries, including Norway and the United Kingdom, reveal significant disparities in access to legal aid [3], [4]. For instance, in Norway, only approximately 9% of the population qualifies for legal aid [3], while in the UK, issues such as means testing and asset thresholds further exacerbate the accessibility challenges [4].

The means testing process, commonly used to assess eligibility for legal aid, often overlooks crucial factors such as homeownership, leading to unjust exclusions from legal assistance [4]. Consequently, individuals with modest incomes or minimal assets may find themselves unable to afford legal representation, further exacerbating socio-economic disparities in accessing justice.

Through rigorous empirical investigation and experimentation, this research endeavor aims to contribute to the advancement of automated legal document classification techniques. By enhancing the accessibility and efficiency of legal aid provision, particularly in identifying human rights violations, this project endeavors to promote equitable access to justice and uphold the integrity of legal systems.

II. METHODLOGY

A. Proposed work:

The proposed work aims to develop a legal documentation classifier capable of handling long documents using a sliding window approach for extended sequence length. Leveraging the ECHR dataset, the system employs a combination of conventional machine learning techniques (SVM[5], DT[6], NB[7], AdaBoost[8]) and transformer-based models (BERT[9], Legal-BERT[10], RoBERTa, BigBird, ELECTRA, XLNet) for classification, with performance evaluation based on accuracy, precision, recall, and F1-score.

As an extension, ensemble methods like the Voting Classifier (RF+AdaBoost) and deep learning models such as LSTM and LSTM+GRU will be integrated to enhance prediction accuracy and model robustness. Furthermore, a user-friendly front-end interface will be developed using Flask, facilitating seamless interaction with the legal documentation classifier. Authentication features will also be incorporated to ensure secure access to the system.

B. System Architecture:

The system begins by inputting data from the ECHR dataset, comprising legal documents. These documents undergo tokenization and vectorization, converting them into numerical representations suitable for machine learning algorithms. Feature selection techniques are applied to identify the most relevant features for classification. The dataset is then divided into a train set and a test set for model training and evaluation. The training module utilizes various algorithms to learn patterns from the data, while the testing module assesses the model's performance on unseen data. Performance evaluation metrics such as accuracy, precision, recall, and F1-score are computed. Identified cases are categorized as non-violated or violated based on the model's predictions. The system architecture incorporates these steps into a cohesive framework, ensuring efficient processing of legal documents and accurate classification results for human rights violation cases.

C. Dataset collection:

The dataset for classifying European Court of Human Rights (ECHR) cases comprises a collection of legal documents obtained from official court records and publicly available sources. These documents encompass a diverse range of cases involving human rights violations, adjudicated by the ECHR. The dataset includes textual information such as case descriptions, legal arguments, court decisions, and relevant contextual details.

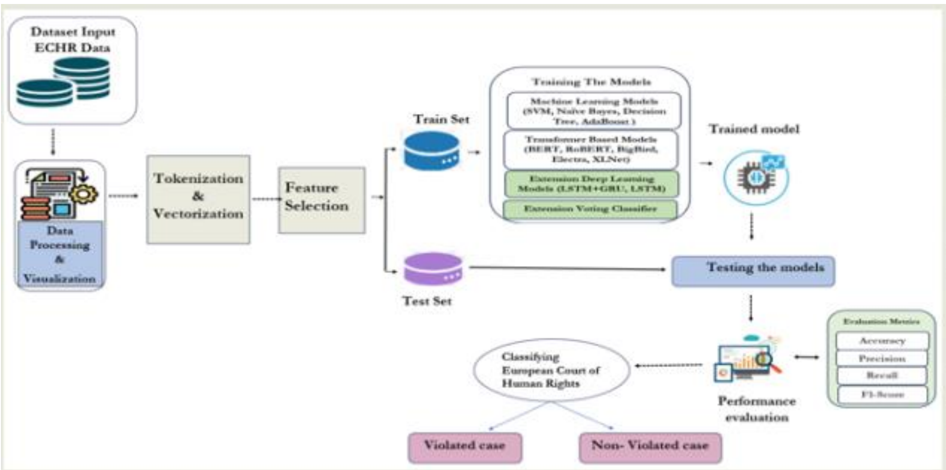

Fig. 1 Proposed Architecture for ECHR Classification using Hybrid Techniques

D. ALGORITHMS:

SVM: Support Vector Machine (SVM)[5] is a supervised machine learning algorithm used for classification and regression tasks. In SVM, data points are plotted in n-dimensional space (where n is the number of features), and a hyperplane is constructed to separate different classes with the maximum margin. In the project of classifying European Court of Human Rights (ECHR) cases using transformer-based techniques, SVM[5] can be employed as a baseline classification algorithm. By training SVM on features extracted from the legal text data, it can learn to classify ECHR cases into different categories based on patterns in the data. SVM's ability to handle high-dimensional data and its effectiveness in binary and multiclass classification make it a suitable choice for this task.

Decision Tree: Decision Tree[6] is a supervised machine learning algorithm used for both classification and regression tasks. It creates a tree-like structure where each internal node represents a feature, each branch represents a decision based on that feature, and each leaf node represents a class label or regression value.

In the project of classifying European Court of Human Rights (ECHR) cases using transformer-based techniques, Decision Tree can be utilized to classify legal documents based on extracted features. By recursively partitioning the dataset into subsets based on feature values, Decision Tree learns decision rules that help classify ECHR cases into different categories, providing insights into the legal text data.

NaiveBayes: Naive Bayes[7] is a probabilistic machine learning algorithm based on Bayes' theorem with the naive assumption of feature independence. It's particularly useful for text classification tasks. In Naive Bayes, each feature is considered independently, which simplifies the calculation of probabilities.

AdaBoost: AdaBoost,[8] short for Adaptive Boosting, is an ensemble learning algorithm used for classification tasks. It combines multiple weak classifiers to create a strong classifier. In AdaBoost[30], each weak classifier focuses on instances that were misclassified by the previous classifiers, adjusting its weights to improve accuracy iteratively.

BERT: BERT[9] (Bidirectional Encoder Representations from Transformers) is a state-of-the-art transformer-based model designed for natural language processing tasks. It utilizes a deep bidirectional learning approach to pre-train representations of text data, capturing contextual information effectively. In the project focused on classifying European Court of Human Rights (ECHR) cases, BERT is employed as a transformer-based technique for document classification. By fine-tuning pre-trained BERT models on the ECHR dataset, the model learns to classify legal documents based on their content, leveraging the rich contextual embeddings provided by BERT. This allows for accurate categorization of ECHR cases into relevant classes, aiding in legal analysis and decision-making processes.

RoBERTa: RoBERTa[10] (Robustly optimized BERT approach) is a variant of the BERT model that further improves performance by addressing pre-training objectives and fine-tuning techniques. It removes the next sentence prediction task and utilizes dynamic masking patterns during pre-training. In the project, RoBERTa is employed as a transformer-based technique for classifying European Court of Human Rights (ECHR) cases. By fine-tuning pre-trained RoBERTa models on the ECHR dataset, the model learns to classify legal documents effectively. RoBERTa's robustness and enhanced pre-training techniques contribute to improved classification accuracy and generalization, enabling more accurate categorization of ECHR cases and aiding in legal analysis and decision-making processes.

BigBird: BigBird is a transformer-based model designed to handle long-range dependencies more efficiently. It introduces sparse attention mechanisms, enabling it to process sequences of up to 8,000 tokens with fewer computations. In the project, BigBird is utilized as a transformer-based technique for classifying European Court of Human Rights (ECHR) cases. By leveraging its ability to efficiently process long documents, BigBird enhances the classification performance on lengthy legal texts, providing accurate categorization of ECHR cases. Its sparse attention mechanism allows for scalability to larger input sizes, making it suitable for handling the extensive legal documentation commonly found in ECHR cases.

ELECTRA: ELECTRA (Efficiently Learning an Encoder that Classifies Token Replacements Accurately) is a transformer-based model designed to improve efficiency and effectiveness in pretraining for natural language understanding tasks. Unlike traditional masked language models, ELECTRA utilizes a novel pretraining task called replaced token detection, which focuses on identifying replaced tokens rather than predicting masked tokens. In the project, ELECTRA is employed as a transformer-based technique for classifying European Court of Human Rights (ECHR) cases. By leveraging its efficient pretraining approach, ELECTRA [34] enhances the model's ability to understand and classify legal documents accurately, contributing to improved performance in ECHR case classification tasks.

XLNet: XLNet is a transformer-based language model that incorporates bidirectional context into pretraining while maintaining autoregressive models' advantages. It leverages permutation-based training, enabling it to consider all possible permutations of the input sequence during training, thereby capturing bidirectional context effectively. In the project, XLNet is utilized as a powerful transformer-based technique for classifying European Court of Human Rights (ECHR) cases. By leveraging its advanced pretraining mechanism and bidirectional context modeling, XLNet enhances the model's ability to understand complex legal texts and classify them accurately, contributing to improved performance in ECHR case classification tasks.

LSTM: Long Short-Term Memory (LSTM) is a type of recurrent neural network (RNN) architecture capable of learning long-term dependencies in sequential data. In the project, LSTM is employed as a deep learning technique for classifying European Court of Human Rights (ECHR) cases. By processing the textual data sequentially and retaining information over long sequences, LSTM enhances the model's ability to capture complex patterns and dependencies within legal documents. Its recurrent nature allows it to effectively model temporal relationships, making it suitable for tasks requiring understanding and analysis of sequential data, such as natural language processing tasks like document classification.

LSTM+GRU: LSTM+GRU combines the strengths of Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), two types of recurrent neural network (RNN) architectures. In the project, LSTM+GRU is utilized as a deep learning technique for classifying European Court of Human Rights (ECHR) cases. By leveraging both LSTM and GRU layers, the model can effectively capture long-term dependencies and temporal patterns within textual data. This hybrid architecture enhances the model's ability to understand and analyze sequential data, improving its performance in tasks such as natural language processing and document classification. LSTM+GRU's integration of LSTM's memory cell structure with GRU's gating mechanisms enables efficient information retention and processing, making it suitable for handling complex sequential data.

III. EXPERIMENTAL RESULTS

Accuracy: The accuracy of a test is its ability to differentiate the patient and healthy cases correctly. To estimate the accuracy of a test, we should calculate the proportion of true positive and true negative in all evaluated cases. Mathematically, this can be stated as:

Accuracy = TP + TN/ TP + TN + FP + FN.

Precision: Precision evaluates the fraction of correctly classified instances or samples among the ones classified as positives. Thus, the formula to calculate the precision is given by:

Precision = True positives/ (True positives + False positives) = TP/(TP + FP)

Recall: Recall is a metric in machine learning that measures the ability of a model to identify all relevant instances of a particular class. It is the ratio of correctly predicted positive observations to the total actual positives, providing insights into a model's completeness in capturing instances of a given class.

Recall = True positives/ (True positives + False Negatives) = TP/(TP + FN)

F1-Score: F1 score is a machine learning evaluation metric that measures a model's accuracy. It combines the precision and recall scores of a model. The accuracy metric computes how many times a model made a correct prediction across the entire dataset.

ML Model	Accuracy	Precision	Recall	f1_score
SVM	0.683	0.563	0.683	0.784
Decision Tree	0.708	0.707	0.708	0.706
NaiveBayes	0.667	0.677	0.667	0.717
AdaBoost	0.700	0.700	0.700	0.700
Extension Voting Classifier	1.000	1.000	1.000	1.000
BERT	0.986	0.986	0.986	0.986
RoBERTa	0.633	0.628	0.628	0.628
BigBird	0.630	0.625	0.625	0.625
ELECTRA	0.807	0.808	0.808	0.808
XLNet	0.560	0.564	0.564	0.564
Extension LSTM	0.927	0.927	0.927	0.927
Extension LSTM+GRU	0.637	0.629	0.682	0.603

Fig.2 Performance Evaluation for all algorithms

IV. CONCLUSION

In conclusion, this study presents a comprehensive approach to legal document classification, leveraging both transformer-based models and traditional machine learning techniques. Through extensive evaluation of models such as BERT, RoBERTa, SVM, and Naive Bayes, valuable insights into their performance are gained, highlighting the potential for transformer-based models in legal text analysis. Furthermore, the efficacy of the extension Voting classifier algorithm as an ensemble method underscores its practical utility, validated through frontend interface testing. By automating document analysis, this research contributes to enhancing accessibility to legal aid, potentially reducing costs and streamlining processes within the legal system. Moving forward, there is ample opportunity for further exploration of advanced transformer models and integration of additional features to enhance accuracy and robustness in legal document classification, laying the groundwork for future research in this domain.

V. REFERENCES

1. J.E.Carlin andJ. Howard,‘‘Legal representation and class justice,’’ UCLA L. Rev., vol. 12, p. 381, Jan. 1964.
2. H. Donald, ‘‘Establishing a national standard: The role of the UK’s minimum income standard in policy and practice,’’ in Minimum Income Standards Reference Budgets. Bristol, U.K.: Policy Press, 2020, pp. 307–318.
3. A.K.ChaturvediandR.Koul,‘‘Legalaidandlegislativeinitiatives,’’ Think India J., vol. 22, no. 14, pp. 11256–11266, 2019.
4. H. Henrik. (2022). Classifying European Court of Human Rights Cases Using Transformer Based Models.[Online].Available:https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/3033966
5. N. Cristianini and J. Shawe-Taylor, An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods. Cambridge, U.K.: Cambridge Univ. Press, 2000.
6. S. B. Kotsiantis, ‘‘Decision trees: A recent overview,’’ Artif. Intell. Rev., vol. 39, no. 4, pp. 261–283, 2013.
7. K. M. Leung, ‘‘Naive Bayesian classifier,’’ Polytech. Univ. Dept. Comput. Sci./Finance Risk Eng., vol. 2007, pp. 123–156, Nov. 2007.
8. R. E. Schapire, ‘‘Explaining adaboost,’’ in Empirical Inference. Cham, Switzerland: Springer, 2013, pp. 37–52.
9. Adhikari, A. Ram, R. Tang, and J. Lin, ‘‘DocBERT: BERT for document classification,’’ 2019, arXiv:1904.08398.
10. Chalkidis, M. Fergadiotis, P. Malakasiotis, N. Aletras, and I.Androutsopoulos, ‘‘LEGAL-BERT: The muppets straight out of law school,’’ 2020,arXiv:2010.02559.